IJSRM
INTERNATIONAL JOURNAL OF SCIENCE AND RESEARCH METHODOLOGY
An Official Publication of Human Journals

Human Journals
Review Article
May 2022 Vol.:21, Issue:3

Is The Normal Distribution A PUN?

IJSRM
INTERNATIONAL JOURNAL OF SCIENCE AND RESEARCH METHODOLOGY
An Official Publication of Human Journals

Kent Olson

Ph.D., Philosophy of Science

Oxford OUDCE Philosophical Society, Scotland

Submitted: 20 April 2022
Accepted: 26 April 2022
Published: 30 May 2022

HUMAN JOURNALS

www.ijsrm.humanjournals.com

Keywords: induction, Hume, normal distribution, scientific method, statistical inference, philosophy of science, Gaussian distribution, Bayes' theorem, Bayes

ABSTRACT

If we act in accordance with a rule, do we know whether or not we are acting rationally? With his famous billiard ball example, Hume pointed out there is nothing in perception that can account for the phenomenon of one billiard hitting another into a pocket, although we expect it to be the case. Statistical inference, which is much more sophisticated than Baconian simple enumerative induction, relies upon "models". These aid us in visualizing "patterns" we see in "nature", for the most part. We have the Gaussian distribution, binomial distribution, the chi-square, and so forth. Bayes' theorem offers contemporary inductivists two great benefits to counter a charge that we are creating inferential patterns out of dust—PUNs, to put it bluntly. In the greater scope of probability theory, it may be argued that parametric tests and frequentist methods fall into Hume's trap by assuming certain distributions exist in nature. This is a short paper arguing that subjectivist Bayesian epistemology avoids skepticism when viewed through the right lens.

I. Is Induction Rational?

Hume notoriously left philosophy in an awkward predicament. To condemn induction completely doesn't seem to be in step with the current philosophical literature. This was Popper's route,[1] and it was met with stern criticism. For example: "most scientific evidence does not bear the logical relationship to theories Popper envisaged" (Howson & Urbach, 2006, 4). Even with the use of syllogistic logic, to have universal generalizations, there needs to be a point where the reasoner must rely upon induction, and it would be irrational to think deduction offers a way out. If Hume is right, the only possible justification for the use of induction would be a principle of uniformity of nature (PUN). In both the *Treatise* and the *Enquiries*, Hume argued as an empiricist that such a principle would need to be informed by observation.

As methods developed, more questions arose. Some have suggested that inductive circles in logic may be benign (Norton, 2021)**.** Furthermore, we may ask if conditionalization with the use of Bayes' theorem is sufficient for our belief in its rationality. If induction has been beneficial, shouldn't its use be continued? Saint-Mont argues that Hume's vague idea of "some uniformity of nature" can be formalized using logical notation (Saint-Mont, 2017, 35). He phrases the main gripe in *A Treatise of Human Nature* as follows:

"All probable arguments are built on the supposition that there is this conformity betwixt the future and the past, therefore can never prove it. This conformity is a matter of fact, and if it must be proved, will admit of no proof but from experience. But our experience in the past can be proof of nothing for the future, but upon a supposition, there is a resemblance betwixt them." (Saint-Mont, 2017, 35)

Pragmatists have many things to say that dovetail nicely with Bayes', surprisingly. We might agree with Peirce that we don't want to accept Hume's skeptical solution of a psychological propensity. He admonishes us that the point of philosophy is to achieve certainty, not to end in doubt (Peirce, 1877, 5). If there is a cash value to our use of a rule, and there is not one when we discontinue, should we not use it? If the answer is *no*, then we ought to continue, and it is rational to do so according to William James' cash-value criterion, which I will hereby accept as a provisional assumption for this paper. If a belief brings us into contact with our desired termini, James would go so far as to call it *true* (James, 1907, 141-155). However, for the time being, let us just assume**:** **i)** that induction *is* used in the sciences; **ii)** that its use has shown cash value; **iii)** if James is right, it is rational to continue to use induction, rather than to discard it. Hume's position is as follows:

"That there is nothing in any object, considered in itself, which can afford us a reason for concluding it; and, that even after the observation of the frequent or constant conjunction of objects, we have no reason to draw any inference concerning and inference beyond those of which we have had experience. [. . .] If reason determined us, it would proceed upon that principle that instances, of which we have had no experience, must resemble those, of which we have had experience, and that the course of nature continues always uniformly the same." (Selby-Bigge, 1888, 152)

II. The Principle and Where it Fits

Keeping in mind **i-iii** above, the onus might be on the inductive skeptic. For us to reason from past experiences to future experiences, Hume argued we needed some sort of principle (PUN) to support the practice. In statistics, the same may hold for inferences from a sample to a population. To be sure, he does offer some way out in saying that induction is a psychological propensity, although we must keep in mind Hume's stance is counter to Bacon's in the *Novum Organum*. This was the first major explication of the inductive method, published originally in 1620. Bacon argued that this was the only way to discover new sciences:

"The syllogism consists of propositions; propositions of words; words are the signs of notions. If, therefore, the notions (which form the basis of the whole) be confused and carelessly abstracted from things, there is no solidity in the superstructure. Our only hope, then, is in genuine induction." (Bacon, 1902, 8)

The syllogistic method promoted by Aristotle only led to philosophical stagnation, according to Bacon's analysis. Statistical methods are a more sophisticated version of inductive inference. The Gaussian distribution allows statisticians to make certain predictions from past instances. According to the central limit theorem, the closer we get to ∞ on a Bernoulli trial, the better the chance our predictions about a population having a certain characteristic are accurate. Let us take

Adolphe Quetelet's example of a sample size of 5738 Scottish militiamen in representing the greater population (Quetelet, 1846, 400). This leads us to our problem.

The argument sketched above stipulated we needed something in observation attending to our reasonings from cause and effect. If we infer from the sample to the greater population, the probability that the next Scottish militiaman has a certain chest measurement, according to the binomial distribution, will follow a certain pattern, assuming an appropriate sample size. Is it possible for a philosopher, given a margin of error, to judge that probabilistic reasoning has provided consistent, reliable results in the past, as we argued previously with the pragmatists? This would be counter to the skeptical conclusion.

Skepticism holds that we have no reason to think that nature operates in a way such that there is a .5 chance that the next Scottish militiaman I observe will have a chest measurement greater than size *n*. I argue that it is more rational to rely upon an inductive practice than not. What is **(a)** the probability that Hume's assessment that induction is irrational, versus **(b)** the probability that the next raven I see will be black (to use another textbook example)? For us to endorse a fully Humean inductive skepticism, the probability of **(a)** should be higher than **(b)**. Ironically, this is roughly the same type of reasoning he employs against miracles in Section X of *An Enquiry Concerning Human Understanding*, and it directs our attention to the rational status of induction itself. He refers to the principle inherent in this reasoning as *contrariety* (Selby-Biggie, 1748, 578).

III. Statistical Inference and Bayes

That a sceptical conclusion is not entirely warranted should be abundantly clear, and we can approach the argument in various ways. Bayesian reasoning is based on betting behavior. This account contrasts with the frequency interpretation of probability. To see wherein probabilistic concepts may resemble PUNs in general, we ought to take our cue from classical statistics. The ontological commitments of subjective Bayesians remove them from certain Humean concerns raised by PUNs. Classical theories of statistics are often called *frequentist*. Some notable frequentists are Antoine Augustine Cournot, John Stuart Mill, John Venn, von Mises, Reichenbach, and Neyman. The differences between the two views highlight how a change in probabilistic methods may avoid the problem of induction. Bayesians concentrate upon degrees of belief, frequentists upon limiting relative frequency. Keynes, Ramsey, and Jeffreys fall under the Bayesian camp. A parameter is a fixed constant or a random variable. The frequentists require a fixed variable, Bayesian reasoning does not. There are other reasons for adopting the latter as well.

There are a few arguments against Bayesian reasoning as a solution to the problem of induction. The objectivist version of Bayesianism seems to possess a few difficulties. Whilst the subjectivist calculates degrees of belief based on betting behavior, the objective version is putatively considered dogmatic, requiring two extra constraints. Even with a subjectivist version, there are rival accounts on how to deal with the problem of induction. Falsificationism, for example, relies upon corroboration rather than verification. Popper argued that scientific discovery followed the logical pattern of *modus tollens*, which is deductive. "Induction, i.e. inference based on many observations, is a myth. It is neither a psychological fact, nor a fact of ordinary life, nor one of scientific procedure" (Popper, 1963, 53).

Another problem falls across the board. The problem of prior probability, which is widely cited, is a chimera, and there is more than one. John D. Norton points out a unique problem for Bayes in that if H is set to 0, then no amount of evidence E can help us to reach a statistic. It might be conjectured whether or not H being 0 as a hypothesis could have anything going for it. Could empirical instances of a phenomena cause an H to appear in a sentient mind? Can something come from nothing? That would be interesting, and thoughts toward a Lockean *tabula rasa* may perhaps spring to mind as kneejerk reaction. Philosophers of mind might have something to say about the argument. Norton writes: “Once $P(H) = 0$, Bayes’ theorem (B) requires $P(H|E) = 0$ for all admissible E. (Conditionalizing on evidence E for which $P(E) = 0$ leads to an undefined posterior $P(H|E)$.) It follows that for any $H' = \sim H$ for which $P(H') = 1$, $P(H'|E) = 1$ as well” (Norton, 2011, 430).

That a prior probability is set to zero is attributed to Popper, and is considered an “historical curiosity” (Zabel, 2011). The obvious answer to this is noninformative priors. Conditionalization is a process wherein the subjective personal degrees of belief of two or more researchers in the field may converge upon the same probability whilst using the theorem.

IV. Parametric and Non-parametric Tests

We have encountered some problems that may infect statistical inferences. The first one was solved by accepting the subjectivist view, the second by grasping the idea that noninformative priors have a minimal overall effect on the data. Bayes' theorem is to be used reiteratively. This is the principle of conditionalization at work, and it should serve as an antidote to potentially epistemically immodest claims. One further tangent I have yet to see exploited in the literature is related to hypothesis testing in statistics. Bayesian inference is usually considered non-parametric. Our final concern is whether or not a normal distribution is a principle of uniformity of nature (Mittel, 2017). This appears to be the case, due to some of the ontological commitments bound up with the method.

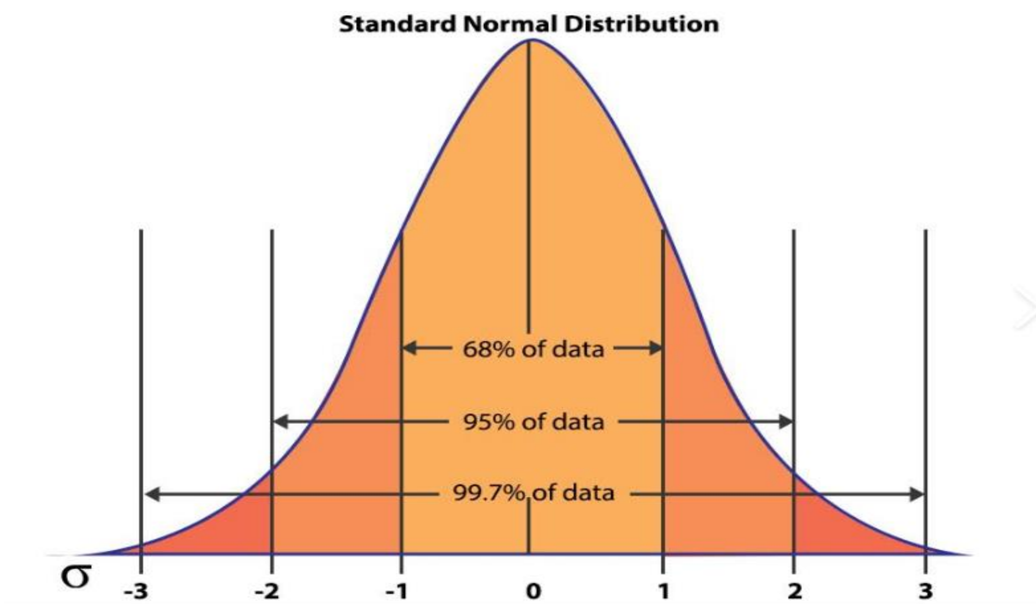

Fig. 1. An illustration of a normal distribution complete with mean and standard deviations. Does this suggest there are uniformities in nature we cannot fully explicate? Quantra. "Standard Normal Distribution," *QuantiInsti*, 2021. <https://quantra.quantinsti.com/glossary/Standard-Normal-Distribution>.

Parametric testing commits scientists to various assumptions. A PUN may be smuggled in due to assumptions used in hypothesis testing. Let us say a scientist wants to calculate via a hypothesis test. The scientist does not know which test to use. He/she must check assumptions. That is, we want to see whether or not our data follows a certain distribution pattern in these scenarios. Usually, parametric tests assume a normal distribution. Non-parametric tests do not (Kozyrkov, 2020). The technique, I believe, to sidestepping the PUN problem and standing up for inductive inference as a practice in science is to minimalize ontological claims.

V. CONCLUSION

When it comes to a blanket Humean objection we just sketched above, Bayesians live to fight another day. As far as a greater concern over whether the use of induction is rational or not, there is more to the story. Is a psychological propensity a logical method? As others in the philosophy of science have pointed out, it is not (O'Neil, 1989). Is inductive inference on Hume's account *methodological* at all? Even Bacon's simple enumerative induction is a more exhaustive account of how science works—especially if we consider crucial tests. Some have even called Bacon's method "eliminative induction" due to its inclusion of them (Schwartz, 2017). Popper held that the crucial test involved an observation, which seems to be the inverse. Ultimately, we might be able to consider induction rational, given a disjunction between it and Hume's scepticism. Is it more rational to think that the next raven I see is black over a brain-in-the-vat scenario?

Can we deem any form of inductive inference deductively valid? This may indeed be a difficult philosophical challenge, although the Bayesians show us that this might be a meaningless star to pin on one's chest. We have seen both that Humean skeptics have problems convincing us that the use of induction is not rational and that a problem of probability does not take conditionalization into account. It looks as if frequentist interpretations of probability have more trouble than subjectivist Bayesianism, and parametric methods are generally epistemically immodest. At this point, it looks as if the Humean sceptic he has to show something wrong with personal degrees of belief.

ACKNOWLEDGEMENT

The author would like to thank his former teacher Professor Alasdair Richmond of the University of Edinburgh for pointing him in the right direction. The information on statistical inference is vast. *For auld lang syne!*

REFERENCES

1. Bacon, Francis V. 1902. *Novum Organum*. New York: Collier and Son.
2. Bandyopadhyay, Prasanta, and Malcolm Forster (eds.). 2011. *Philosophy of Statistics.* Amsterdam: Elsevier.
3. Cassie Kozyrkov. “Statistics: Are you Bayesian or Frequentist? Towards Data Science,” YouTube Video, 7:02, Sep 2, 2020, <https://towardsdatascience.com/statistics-are-you-bayesian-or-frequentist-4943f953f21b>.
4. Cournot, A.A. 2013. *Exposition of the Theory of Chances and Probabilities*, Tr. Oscar Sheynin. <https://arxiv.org/ftp/arxiv/papers/1902/1902.02781.pdf>.
5. Devansh, Mittal 2019. “The Bell-Shaped Curve: A Common Pattern in Nature,” *Intuitive Physics*. <https://medium.com/intuitive-physics/the-bell-shaped-curve-a-common-pattern-in-nature-85ffc28a8565>.
6. Howson, Colin and Peter Urbach. 2006. *Scientific Reasoning: The Bayesian Approach* 3rd ed. Open Court: Chicago and LA Salle.
7. Hume, D. 1888. *A Treatise of Human Nature.* Selby-Bigge (e.d.). Oxford: Clarendon Press.
8. James, W. 1907. “Pragmatism’s Conception of Truth,” *The Journal of Philosophy, Psychology*, *and Scientific Methods* Vol. 4, No. 6: 141-155.
9. Norton, John D. 2011. “Challenges to Bayesian Confirmation Theory,” *Philosophy of Statistics*, Bandyopadhyay, Prasanta, and Malcolm F.(eds.). Amsterdam: Elseveir: 391-439.
10. O’Neil, J. 1989. “Two Problems of Induction” The British Journal of the Philosophy of Science. Vol.40, No.1:121-125.
11. Peirce, Charles S. 1877. “The Fixation of Belief,” In *Popular Science Monthly* 12.
12. Popper, Karl P. 1935. *The Logic of Scientific Discovery*. London and New York: Routledge.
13. Popper, Karl R. 1963. *Conjectures and Refutations*. London: Routledge.
14. Quantra.2021.“Standard Normal Distribution,” *Quantinsti*. <https://quantra.quantinsti.com/glossary/standars-normal-distribution>.
15. Quetelet, Adolphe L. 1846. *Lettres a S.A.R. le Duc Regnant de Saxe-Cobourg et Gotha, sur la Theorie des Probabilites, Appliquee aux Sciences Morales et Politiques*. Brussels: M. Hayes.
16. Saint-Mont, U. 2017. “Induction: The Glory of Science and Philosophy,” *Phil SciArchive*. Nordhousen: University of Applied Sciences.
17. Saint-Mont, U. 2020. *Induction: A Logical Analysis*, *Foundations of Science Foundations of Science*. <https://doi.org/10.1007/s10699-020-09683-z>.
18. Schwartz, D. 2017. “Crucial Instances and Bacon’s Quest for Certainty,” *University of Chicago Press*. Vol.7, No.1.
19. Selby-Bigge, L.A. (ed.). 2009. *David Hume, An Enquiry Concerning Human Understanding, 1748*. (577-586) <http://philosophyfaculty.ucsd.edu/faculty/ewatkins/HUM4Texts/Hume-Miracles%2810%29.pdf>.
20. Van Fraassen, Bas C. 1988. “The Problem of Old Evidence,” *Philosophical Analysis*. 19:153-165.
21. Zabell, Sandy L. 2011. “Carnap and the Logic Inductive Inference,” *Handbook of the History of Logic*. Elsevier, Vol.10: 265-309.